

Homework 2.2 Question

Write a program in the language of your choice that will remove the grade of type "homework" with the lowest score for each student from the dataset in the handout. Since each document is one grade, it should remove one document per student. This will use the same data set as the last problem, but if you don't have it, you can download and re-import.

The dataset contains 4 scores each for 200 students.

First, let's confirm your data is intact; the number of documents should be 800.

```
use students
db.grades.count()
```

Hint/spoiler: If you select homework grade-documents, sort by student and then by score, you can iterate through and find the lowest score for each student by noticing a change in student id. As you notice that change of student_id, remove the document.

To confirm you are on the right track, here are some queries to run after you process the data and put it into the *grades* collection:

Let us count the number of grades we have:

```
db.grades.count()
```

The result should be 600. Now let us find the student who holds the 101st best grade across all grades:

```
db.grades.find().sort( { 'score' : -1 } ).skip( 100 ).limit( 1 )
```

The correct result will be:

```
{ "_id" : ObjectId("50906d7fa3c412bb040eb709"), "student_id" : 100,
  "type" : "homework", "score" : 88.50425479139126 }
```

Now let us sort the students by `student_id`, and `score`, while also displaying the `type` to then see what the top five docs are:

```
db.grades.find( { }, { 'student_id' : 1, 'type' : 1, 'score' : 1, '_id' :
0 } ).sort( { 'student_id' : 1, 'score' : 1, } ).limit( 5 )
```

The result set should be:

```
{ "student_id" : 0, "type" : "quiz", "score" : 31.95004496742112 }  
{ "student_id" : 0, "type" : "exam", "score" : 54.6535436362647 }  
{ "student_id" : 0, "type" : "homework", "score" : 63.98402553675503 }  
{ "student_id" : 1, "type" : "homework", "score" : 44.31667452616328 }  
{ "student_id" : 1, "type" : "exam", "score" : 74.20010837299897 }
```

To verify that you have completed this task correctly, provide the identity of the student with the highest average in the class with following query that uses the aggregation framework. The answer will appear in the `_id` field of the resulting document.

```
db.grades.aggregate( { '$group' : { '_id' : '$student_id', 'average' : {  
$avg : '$score' } } }, { '$sort' : { 'average' : -1 } }, { '$limit' : 1 }  
)
```

Enter the student ID below. Please enter just the number, with no spaces, commas or other characters.